

KOOKKURRENZANALYSE KONTRASTIV

Zum Nutzen von Korpusanalysemethoden für die bilinguale lexikographische Praxis – am Beispiel des GDTAW

Kathrin Steyer – Marie Vachková

Die folgenden Ausführungen knüpfen unmittelbar an den Artikel von Cyril Belica und Kathrin Steyer in diesem Band an und zeigen, wie die dort beschriebenen methodischen Prämissen der Korpusanalyse, speziell der am IDS entwickelten statistischen Kookkurrenzanalyse (KA)¹, auf ein konkretes Anwendungsgebiet hin adaptiert und spezifiziert werden können. Die bei der Erstellung von Wörterbuchartikeln (WBA) im GDTAW untersuchten Wortverbindungen werden aus den Resultaten der Kookkurrenzanalyse gewonnen, indem sie unter kontrastivem Gesichtspunkt gesichtet werden. Das angestrebte Ziel ist – unter der Berücksichtigung des vorgesehenen Adressatenkreises und seiner Bedürfnisse (vgl. den einleitenden Aufsatz in diesem Sammelband) – die tschechischen Übersetzungsmöglichkeiten zu erforschen, um dem Wörterbuchbenutzer eine möglichst umfassende Information über den Bedeutungsumfang des deutschen Lexems und eine repräsentative Auswahl an deutschen Verwendungsstrukturen zu bieten. Es soll auf den diesen Strukturen inhärenten Wortschatzreichtum hingewiesen werden, der sich u.a. in dem Angebot von benutzbaren Synonymen in der Zielsprache Tschechisch niederschlagen kann. So ist z.B. (wie zwischen jedem lexikografisch konfrontierten Sprachenpaar) die Asymmetrie im Bereich der Polysemie darzustellen, und zwar nicht nur – wie bisher zumeist geschehen – auf der *langue*-Ebene, indem einem deutschen Lexem (hier als Beispiel dem Adjektiv *frei*) die tschechischen Systemäquivalente *volný*, *svobodný* gegenübergestellt werden können, sondern vor allem auf der Ebene der *parole*, d.h. kontextgebunden. Korpusanalytische Untersuchungen wie die Kookkurrenzanalyse bieten durch ihre genuin syntagmatische Perspektive neue Möglichkeiten zur Erfassung kontextgebundener Strukturen im Kontrast. Wenn die Verwendungsstrukturen der bearbeiteten Lexeme nicht wörterbuchunabhängig – wie in einer deskriptiven Studie – untersucht werden, sondern vor allem im Hinblick auf die Darstellung des jeweiligen Lexems für den nicht-muttersprachlichen Wörterbuchbenutzer, sind neben den Kookkurrenzlisten in Form von Clusterhierarchien und syntagmatischen Mustern vor allem auch die dazugehörigen KWIC-Übersichten (*Keyword-in-Context* = Kontextzeilen des Suchwortes) unverzichtbare Ergänzungen. Sie werden als „Belegsammlung“ verstanden und benutzt, da sich die statistisch ermittelten Kookkurrenzstrukturen für den Nicht-

¹ CYRIL BELICA: Statistische Kollokationsanalyse und Clustering. COSMAS – Korpusanalysemodul. © 1995 Institut für Deutsche Sprache, Mannheim.

muttersprachler oft nicht problemlos erschließen. Das Hauptkriterium muss das Erreichen einer möglichst treffenden und nachvollziehbaren Darstellung der Strukturbedeutung des Wortes und seiner signifikanten Satelliten in typischen Verwendungen sein. Auch periphere Verbindungen sind von Interesse, um dem Ziel des Wörterbuchs, ein Nachschlagewerk vor allem auch für Fortgeschrittene zu sein, gerecht zu werden. Dabei wird stets ein Vergleich mit den gängigen einsprachigen und zweisprachigen Nachschlagewerken angestellt. In den folgenden Ausführungen soll skizziert werden, inwieweit sich die Erschließung der Korpora mittels der KA als eine ergiebige Hilfe für diese Zielsetzung erweisen kann. Es geht im Kern darum herauszuarbeiten, worin der Fortschritt durch die neuen empirischen Methoden hauptsächlich besteht. Mit dem Argument, dass diese Methoden zur Erfassung eines großen Spektrums an Kontext(-belegen) führen, ist es allein nicht getan. Auch in der nicht korpusbasierten Lexikografie war und ist die Darstellung der Übersetzungssynonyme auf die Beschreibung der Lexeme in Minimalkontexten konzentriert. Bei den stärker benutzerorientierten Nachschlagewerken waren dabei zwei Herangehensweisen gängige Praxis: Entweder wurde der Wörterbuchartikel entsprechend der semantischen Struktur eines deutschen Bedeutungswörterbuchs mit Teilbedeutungen oder Lesarten angelegt. Der Nachteil bestand jedoch darin, dass die syntaktischen Verwendungsmuster in diverse Teilbedeutungen zerstreut dargestellt wurden. Oder man hat den WBA auf der tschechischen Seite nach den möglichen Übersetzungsäquivalenten gegliedert, was einer Reduzierung des Bedeutungsumfangs und des Spektrums der Beziehungen zwischen den einzelnen Teilbedeutungen in der Ausgangssprache Deutsch gleichkam. Die Bedeutungsstruktur des Quellenwörterbuchs ging wegen einer mehr oder weniger asymmetrisch entwickelten Polysemie verloren. Aus diesem Grund griffen die fortgeschritteneren Benutzer lieber gleich zu einem einsprachigen deutschen Wörterbuch. Als dritter möglicher Weg blieb die Kombination von beiden Verfahren, allerdings mit sehr unterschiedlichen Ergebnissen. Die ausgewählten Beispiele dienten als Belege für die aufgeführten Äquivalente. In manchen Fällen boten sie dem Benutzer darüber hinaus auch andere, z.B. syntaktisch oder lexikalisch interessante Übersetzungsvarianten an. Es ist nun zu zeigen, dass die Kookkurrenzanalyse vor allem bei den Wortarten, die im deutschen Satz eine ausgeprägte syntaktische Funktion innehaben (Adjektive, Verben), für eine benutzerfreundliche Strukturierung eines WBA von großem Nutzen ist. (Vgl. die auf den syntaktischen Strukturen basierende Strukturierung der WBA im Langenscheidts Großwörterbuch Deutsch als Fremdsprache). Deshalb haben wir uns für eine weitere Analyse des Adjektivs *frei* entschieden. Der Eintrag zu *frei* im neu aufgelegten und ergänzten Deutsch-tschechischen Wörterbuch von H. Siebenschein et al. (weiter nur Sieb) (Leda 2002) stellt eine solche Kombination im oben beschriebenen Sinne dar, was anhand des Wörterbucheintrages verdeutlicht werden

soll. Die Linearität der ursprünglichen grafischen Ausgestaltung in Sieb wurde von den Autorinnen des vorliegenden Beitrages aufgehoben, da die Anordnung der Teilbedeutungen untereinander eine übersichtlichere Platzierung der hinzugefügten Kommentare in Blockschrift erlaubt. Diese sind vor allem für den deutschen Leser gedacht, aber auch für den Vergleich des vorhandenen und des einzubindenden Wortmaterials:

Wörterbucheintrag zu *frei* in Sieb (2002, Bd. 1, S. 550)

I.

1. *svobodný, nezávislý, volný*; –e Künste, Wissenschaften, Berufe *svobodná umění, svobodné vědy, svobodná povolání*
2. *svobodný (neženaný, neprovdaný)*; [OHNE BEISPIELANGABE]
3. *nenucený, volný, smělý* (chování, řeč) 1 [KOLLOKATOR ANGEGEBEN] (Benehmen, Rede), [OHNE BEISPIELANGABE]
4. *prázdný, volný* (čas); wie verbringen Sie Ihre –e Zeit?
5. *přímý, upřímný*; [OHNE BEISPIELANGABE]
6. *volný, bezplatný* [OHNE BEISPIELANGABE]
7. *dobrovolný* (rozhodnutí); 1 [KOLLOKATOR ANGEGEBEN] (Entscheidung), [OHNE BEISPIELANGABE]
8. *svobodný, volný, neomezený* (obchod); 1 [KOLLOKATOR ANGEGEBEN] (Handel), [OHNE BEISPIELANGABE]
9. *volný, prázdný, nepronajatý* (byt); 1 [KOLLOKATOR ANGEGEBEN] (Wohnung), [OHNE BEISPIELANGABE]
10. f. von etw. *prostý čeho, bez čeho*; f. von Furcht, von Haß, von Leidenschaft *bez strachu, nenávisti, vášně*;

II. jmenné vazby (= nominale Verbindungen): (IN ALPHABETISCHER ABFOLGE)

aus –em Antrieb, aus –em Willen, aus –em z *vlastního popudu, z vlastní vůle, o své újmě*;

–es Benehmen 1. *volné, nenucené chování*; 2. *nevázané chování*;

e –e Bühne *osvobozené divadlo*; –e Fahrt genießen *mít jízdu zdarma*;

fyz. r –e Fall *volný pád*;

–es Feld *širé pole*;

j-n auf den –en Fuß setzen *propustit koho na svobodu*;

–e Hand haben *mít volnou ruku, volné pole*;

j-m –e Hand geben, lassen *ponechat komu volnost, dát komu plnou moc*;

obch. F. ins Haus *franko dům*;

unter –em Himmel *pod širým nebem*;

e sieben –en Künste *sedm svobodných umění*;
 einer Sache –en Lauf lassen *dá(va)t něčemu volný průběh, nebránit čemu*;
 –e Liebe *volná láska*;
 –e Medikamente *léky volně k dostání, bez receptu*;
 –es Meer *širé moře*;
 –es Spiel haben *mít volnou ruku, volné pole*;
 eine –e Stelle *volné, uvolněné místo*;
 r Zug blieb auf freier Strecke stehen *vlak se zastavil na trati v polích*;
 ist r Sessel frei? *Je ta židle volná?*;
 etw. aus –en Stücken tun *udělat co z vlastního popudu, rozhodnutí, sám od sebe*;
 morgen habe ich einen freien Tag *zítra mám volno*;
 –e Wahl haben *mít svobodnou volbu, mít na vybranou*;
 s Zimmer ist f. *pokoj je volný, nepronajatý*;

III. slovesné vazby:

diesmal ist er noch f. ausgegangen *tentokrát ještě vyvázl bez trestu*
 f. haben *mít kdy, mít volno, prázdno*
 den Weg f. machen *uvolnit, klesit cestu*
 etwas f. heraus sagen *mluvit přímo, bez obalu*
 f. sein *být volný, nezávislý*
 f. sein von etw. *být prost, být zbaven čeho*
 darf ich so f. sein? *Dovolíte?*
 f. sprechen **1.** *mluvit volně, bez obalu*; **2.** *mluvit z paměti, spatra*

Das Informationsangebot des Wörterbuchartikels *frei* in Sieb blieb, wie viele andere WBA auch, in der Neuauflage unverändert. Einzige Ausnahme: Die Benennung für die DDR-Jugendorganisation *Freie Deutsche Jugend* wurde gestrichen. Insgesamt aber wird die Konzeption leider nicht angetastet.

Im Folgenden wird der WBA „frei“ mit den KA-basierten Ergebnissen verglichen, um eine effektive Gewinnung von Wortmaterial im Sinne der eingangs erwähnten Desiderate zu demonstrieren. Alle behandelten Wortverbindungsbeispiele basieren auf statistisch signifikanten Kookkurrenzclustern². Zunächst seien

² Zur Beschreibung der Funktionsweise der statistischen Kookkurrenzanalyse vgl. www.ids-mannheim.de/kt/misc/tutorial.html. Zu Möglichkeiten linguistischer Interpretationen und lexikografischer Anwendungen vgl. STEYER u. a. 2000, 2003, 2004b.

die entsprechenden **binären** Kollokationen mit dem Bezugswort *frei*, die diesen ausgewählten Beispielen zu Grunde liegen, aufgelistet:

| Bezugswort | primärer Kookkurrenzpartner |
|------------|-----------------------------|
| FREI | EINTRITT |
| FREI | VERKAUF |
| FREI | FUSS |
| FREI | FAHRT |
| FREI | WILDBAHN |
| FREI | HIMMEL |
| FREI | LAUF |
| FREI | MARKT |
| FREI | WEG |
| FREI | THEATERGRUPPEN |
| FREI | EMPFANGBAR |
| FREI | HAUS |
| FREI | FRANK |
| FREI | STÜCKEN |
| FREI | |

In der nächsten Grafik sind die statistisch ermittelten Cluster vollständig abgebildet (einschließlich weiterer spezifizierender Kookkurrenzpartner), so wie sie von der automatischen KA berechnet wurden. Die Gliederung in die Gruppen a(1)–d(3) verweist auf die Kategorisierungen der nachfolgenden Analyse. Die Cluster werden auf zweifache Weise visualisiert: Auf der linken Seite ist die hierarchische Clusterstruktur und auf der rechten Seite sind die berechneten typischen syntagmatischen („natürlichsprachigen“) Muster abgebildet:³

³ Auf die detaillierte Interpretation dieser Ergebnisse kann an dieser Stelle nicht eingegangen werden. Es bleibt zu betonen, dass es sich in dieser Darstellung um Ergebnisse reiner Berechnungen handelt, die auch zu Resultaten führen können, die aus unserer Sicht zu verwerfen sind. So verbirgt sich hinter die Zeichenkette *pd.* im Cluster *frei – eintritt* in den Zeitungen gehäuft vorkommendes Kürzel, das natürlich für eine linguistische Interpretation bzw. eine lexikografische Bearbeitung irrelevant ist.

a(1)

| | | | | | | | |
|---|----|---|-------|-----------------------|------|------|---|
| + | -2 | 1 | 54160 | Eintritt ist Kollekte | 155 | 96% | Der Eintritt [...] ist [...] frei (es wird eine) Kollekte erhoben |
| + | -2 | 1 | 54160 | Eintritt ist frei.pd | 27 | 96% | Der Eintritt [...] ist frei.pd |
| + | -2 | 1 | 54160 | Eintritt ist | 2929 | 86% | Der Eintritt [...] ist [...] frei |
| + | -2 | 1 | 54160 | Eintritt Kollekte | 197 | 93% | Der Eintritt [ist] frei (es wird eine) Kollekte |
| + | -2 | 1 | 54160 | Eintritt frei.pd | 31 | 100% | Der Eintritt [ist] frei.pd |
| + | -2 | 1 | 54160 | Eintritt | 6227 | 71% | Der Eintritt [ist] frei |

a(2)

| | | | | | | | |
|---|---|---|-------|----------------------------------|------|------|---|
| + | 1 | 1 | 37539 | Verkauf Abonnement Theatergemein | 5 | 100% | Abonnement F grün Theatergemeinde grün freier Verkauf |
| + | 1 | 1 | 37539 | Verkauf Abonnement | 153 | 98% | Abonnement ... freier Verkauf |
| + | 1 | 1 | 37539 | Verkauf Miete | 154 | 98% | Miete ... und freier Verkauf |
| + | 1 | 1 | 37539 | Verkauf Theatergemeinde | 81 | 92% | Uhr Theatergemeinde [...] freier Verkauf |
| + | 1 | 1 | 37539 | Verkauf | 3683 | 92% | freier [...] Verkauf |

b(1)

| | | | | | | | |
|---|---|---|-------|-----------------------|------|------|--|
| | 1 | 1 | 26352 | Fuß gesetzt angezeigt | 2 | 50% | freien Fuß gesetzt ... angezeigt |
| + | 1 | 1 | 26352 | Fuß gesetzt Kautio | 48 | 100% | gegen Kautio auf freien Fuß gesetzt worden |
| + | 1 | 1 | 26352 | Fuß gesetzt | 713 | 99% | wieder auf freien Fuß gesetzt |
| + | 1 | 1 | 26352 | Fuß angezeigt | 556 | 99% | wurde auf freiem Fuß [...] angezeigt |
| + | 1 | 1 | 26352 | Fuß Kautio | 92 | 60% | gegen Kautio auf freien Fuß gesetzt worden |
| + | 1 | 1 | 26352 | Fuß | 2487 | 58% | wieder auf freiem Fuß angezeigt |

b(2)

| | | | | | | | |
|---|---|---|------|--------------|------|-----|------------------------------------|
| + | 1 | 1 | 7338 | Fahrt Bürger | 136 | 41% | Freie freie Fahrt für freie Bürger |
| + | 1 | 1 | 7338 | Fahrt | 1026 | 50% | freie [...] Fahrt |

b (3)

| | | | | | | | |
|---|---|---|------|----------|-----|-----|--------------------|
| + | 1 | 1 | 5023 | Wildbahn | 355 | 76% | in freier Wildbahn |
|---|---|---|------|----------|-----|-----|--------------------|

c (1)

| | | | | | | | |
|---|---|---|-------|--------|------|-----|---------------------------|
| + | 1 | 1 | 10341 | Himmel | 1209 | 92% | unter freiem [...] Himmel |
|---|---|---|-------|--------|------|-----|---------------------------|

c (2)

| | | | | | | | |
|---|---|---|------|-------------------------|------|-----|--|
| + | 1 | 1 | 9171 | Lauf gelassen Phantasie | 13 | 84% | ihrer Phantasie freien Lauf gelassen |
| + | 1 | 1 | 9171 | Lauf gelassen | 96 | 79% | freien Lauf gelassen |
| + | 1 | 1 | 9171 | Lauf lassen Phantasie | 50 | 82% | ihrer Phantasie [...] freien Lauf [...] lassen |
| + | 1 | 1 | 9171 | Lauf lassen | 336 | 88% | ihrer ... freien [...] Lauf [zu] lassen |
| + | 1 | 1 | 9171 | Lauf Phantasie | 96 | 96% | ihrer Phantasie [...] freien Lauf lassen |
| + | 1 | 1 | 9171 | Lauf | 1023 | 95% | freien [...] Lauf lassen |

| | | | | | | | |
|---|---|---|------|------------------|------|-----|--|
| + | 1 | 1 | 5943 | Markt Zugang | 19 | 42% | der Zugang zum zu ... freien Markt |
| + | 1 | 1 | 5943 | Markt Wettbewerb | 10 | 40% | freien Markt im fairen Wettbewerb |
| + | 1 | 1 | 5943 | Markt können | 33 | 57% | auf dem im freien [...] Markt [...] zu] können |
| + | 1 | 1 | 5943 | Markt | 1064 | 71% | auf dem freien [...] Markt |

| | | | | | | | |
|---|----|----|------|-------------|------|------|--|
| + | -5 | -1 | 4886 | Weg machen | 120 | 86% | den Weg [für ...] frei [zu] machen für |
| + | -5 | -1 | 4886 | Weg gemacht | 76 | 100% | den Weg [für ...] frei gemacht für |
| + | -5 | -1 | 4886 | Weg | 1624 | 88% | den der Weg [...] frei für |

| | | | | | | | |
|---|---|---|-----|------------------------|----|-----|---|
| + | 1 | 1 | 699 | Theatergruppen Theater | 5 | 40% | Theater ... die freien [...] Theatergruppen |
| + | 1 | 1 | 699 | Theatergruppen | 71 | 33% | der die freien [...] Theatergruppen in ... |

d (1)

| | | | | | | | |
|---|---|---|------|--------------|----|------|------------------------------------|
| + | 1 | 1 | 1034 | empfangbaren | 74 | 100% | im frei empfangbaren Fernsehen ... |
|---|---|---|------|--------------|----|------|------------------------------------|

d (2)

frei Haus

| | | | | | | | |
|---|---|---|------|------|-----|-----|-----------------|
| + | 1 | 1 | 1236 | Haus | 532 | 71% | frei [...] Haus |
|---|---|---|------|------|-----|-----|-----------------|

d (3)

| | | | | | | | |
|---|----|----|-----|-------|----|-----|------------------|
| + | -2 | -2 | 923 | frank | 69 | 98% | frank [und] frei |
|---|----|----|-----|-------|----|-----|------------------|

d (4)

| | | | | | | | |
|---|---|---|------|---------|-----|-----|--------------------------|
| + | 1 | 1 | 2682 | Stücken | 293 | 97% | aus freien [...] Stücken |
|---|---|---|------|---------|-----|-----|--------------------------|

Die Sichtung des Materials verläuft unter dem kontrastiven Gesichtspunkt auf allen Ebenen, auf denen die Asymmetrien zwischen den verglichenen Sprachen vorkommen und somit als vorzubeugende Fehlerquelle im GDTAW berücksichtigt werden müssen. Die als relevant zu verzeichnenden Kookkurrenzen aus der KA-Übersicht wurden in Gruppen strukturiert. Die Strukturierung des Materials erfolgte nach Kriterien des Wörterbuchs. In der folgenden Übersicht werden jedoch nur die für einzelne Gruppen symptomatischen Beispiele gezeigt, einer eingehenden kontrastiven Betrachtung mit reziproker Überprüfung der Kookkurrenzpartner wird sich ein anderer Aufsatz widmen.

a. Lexikalische Äquivalenz – Asymmetrie in den Ausbaumöglichkeiten syntaktischer Strukturen

Die folgenden zwei Beispiele sollen verdeutlichen, dass eine vorliegende Äquivalenz auf lexikalischer Ebene nicht unbedingt Bestand hat, wenn man diese Relationen in einem weiteren syntaktischen Kontext ausbaut. Es entstehen insofern Asymmetrien, als freien syntaktischen Variationen in der einen Sprache – kontextuelle – Restriktionen in der anderen gegenüber stehen. Hier wird u. E. auch deutlich, dass die Erfassung rein lexikalischer Kollokationsrelationen (z.B. alle nominalen Partner für ein Adjektiv) für die Bedürfnisse der Textproduktion in einer Fremdsprache wenig hilfreich ist.⁴

- 1) Substantiv **Eintritt**: Die zum Kookkurrenzcluster **FREI – EINTRITT** ermittelten typischen syntagmatischen Muster weisen auf zwei syntaktische Verwendungsmuster hin: Der Eintritt ist frei *vstup je volný*, Eintritt frei *vstup volný*. Diese zwei parallel vorkommenden syntaktischen Strukturen sollten verwendungsgetreu inventarisiert werden, und nicht nur eine Strukturvariante, vgl. oben bei Sieb oder vgl. DUW bzw. GWDS).
- 2) Substantiv **Verkauf**: Betrachtet man das Kookkurrenzcluster **FREI – VERKAUF** aus rein kollokativ-binärer Sicht, lässt sich ein tschechisches Äquivalent ohne Weiteres zuordnen: *volný prodej*. Die Interferenzprobleme können aber auf der Ebene der syntaktischen Realisierungen entstehen. Die verbalen Ergänzungen zu *freier Verkauf* zeigen die Vielfalt der Möglichkeiten hinsichtlich symmetrischer (SYM) und asymmetrischer (ASYM) Verwendungsstrukturen, vgl. im freien Verkauf/Handel sein vs. *být volně na prodej*, *volně prodejný*. Alle vorgefundenen Ausdrucksmöglichkeiten werden in der Wörterbuch-Datenbank abgespeichert, nicht alle gehen jedoch in den WBA ein, vgl. SYM Die Karten gehen in den freien Verkauf *Lístky jdou do volného prodeje*.

⁴ Vgl. dazu Hausmann, der Kollokationen zwar nach wie vor als grundsätzlich binäre Einheiten ansieht, aber einräumt, dass sich zwei Kollokationen zu einer Tripel-Struktur verbinden können und als solche üblich sind (HAUSMANN 2004, S. 316). Nach unserer Auffassung handelt es sich in den allermeisten Fällen um multiple Cluster (BELICA – STEYER in diesem Band; STEYER 2004a, S. 97).

ASYM Die Karten stehen zum freien Verkauf. *Lístky jsou volně na prodej, jsou ve volném prodeji.*

SYM den Rest der Karten in den freien Verkauf geben *dát zbytek lístků do volného prodeje.* ASYM In der Tageskasse liegen 5 000 Karten zum freien Verkauf *U denní pokladny je na prodej 5 000 lístků.*

Die denkbare syntaktische Alternative **der Verkauf ist frei* wird durch die Kookkurrenzanalyse nicht gestützt. Die attributive Stellung von *frei* ist in genau diesem Verwendungskontext fix: ‚Der Verkauf (zumeist von Eintrittskarten) ist nicht eingeschränkt/begrenzt (z.B. durch Abonnements, geschlossene Veranstaltungen)‘. Es handelt sich um einen domänen gebundenen formelhaften Gebrauch, der auch so zu kennzeichnen ist. Im Tschechischen ist jedoch eine prädikative Verwendung *Nyní je (jejich) prodej volný* aufgrund der typologisch bedingten freien Wortfolge möglich.

b. Pragma-semantische Äquivalenz – strukturelle Asymmetrie (lexikalisch u/o syntaktisch)

- 1) Die Wortverbindung **auf freien Fuß setzen** hat im Tschechischen die Entsprechung *být propuštěn na svobodu* (= frei gelassen werden, wörtlich: in die Freiheit entlassen sein), wobei das syntagmatische Kookkurrenzmuster *gegen Kautions auf freien Fuß gesetzt* (s.o.) noch eine weitere syntaktische Einbettung verdeutlicht, die es im Wörterbuch zu berücksichtigen gilt. Das syntagmatische Kookkurrenzmuster **auf freiem Fuß angezeigt werden** (= auf frischer Tat; unmittelbar am Tatort eine Anzeige der Polizei erhalten) ist nur in den österreichischen IDS-Korpora belegt (auch das sollte gekennzeichnet werden) und würde nur mit *být chyčen na svobodě* umschrieben werden können. Das Beispiel zeigt sehr deutlich, dass binäre Kollokationsrelationen wie FREI – FUSS zwar in gewisser Weise ein Indiz für konzeptuelle Zusammenhänge im Lexikon sein können, dass aber – vor allem für die Bedürfnisse der fremdsprachlichen Textproduktion – erst die syntaktischen Realisierungen wirklichen Aufschluss über die Art und Weise der Kombination der Elemente bringen und dass es oft sogar innerhalb der Subcluster zu semantischen und pragmatischen Ausdifferenzierungen kommt, die es zu berücksichtigen gilt.
- 2) Die Wortverbindung **freie Fahrt [bekommen]** wird bei SIEB nur als *dostat volnou ruku* (wörtlich *freie Hand bekommen*), *svobodu rozhodovat* (wörtlich *Entscheidungsfreiheit erhalten*) interpretiert. Die im Korpus sehr prominente Bedeutung, ohne Hindernisse fahren (z.B. auf der Autobahn nach einem Stau/Unfall, ohne Mautgebühren usw.) fehlt. Die ergänzenden statistisch signifikanten Verben (z.B. **freie Fahrt** → haben – bekommen – gewähren – garantieren – signalisieren – verschaffen – ermöglichen – geben usw.) sind in ihrer disambiguierenden Funktion in diversen Kontexten weiter zu untersuchen und im

WBA anzuführen. Jede Isolierung einer speziellen Wortverbindung (wie in SIEB) erscheint aus der Sicht des Benutzers problematisch, vor allem im Hinblick auf die Funktion des Wörterbuchs als Übersetzungshilfe und Versicherungsinstanz zugleich.

- 3) Die Wortverbindung *freie Wildbahn* (in deutsch-tschechischen Wörterbüchern gar nicht inventarisiert!) kann im Tschechischen nur umschrieben werden, etwa durch *v divoké přírodě* – je nach Kontext. Im DUW ist die Wendung *in freier Wildbahn* zwar vertreten, wird aber nur als ein gebundenes Lexem präsentiert (**Wild|bahn**, die: in der Fügung **freie W.** (*freie Natur*): Tiere in freier W. beobachten). Das ist insofern legitim, als auch die KA das Adjektiv *frei* als den mit Abstand signifikantesten Partner von *Wildbahn* ausweist. Die KA zeigt auch Verwendungen ohne das Adjektiv *frei* (ca. 100 Belege). Diese sind jedoch sehr heterogener Art; es lässt sich noch kein einheitliches Verwendungsmuster des freien Lexems *Wildbahn* erkennen. Hier scheint ein Usualisierungsprozess im Gang zu sein, der eine ‚Deunikalisierung‘ und damit Lexikalisierung der Komponente *Wildbahn* zur Folge haben könnte. Dies bedarf jedoch eingehender Analysen. Für die in das Wörterbuch aufzunehmenden Verwendungen sind Muttersprachler zu konsultieren.

c. strukturähnliche bis -identische Äquivalenz – kontextuelle Varianz

- 1) Die Wortverbindung *unter freiem Himmel* ist in syntaktischer Hinsicht voll-äquivalent (Präpositionalphrasen) und in lexikalischer Hinsicht teiläquivalent (*unter* → *pod*, *Himmel* → *nebe*, *širý* – *weit*). (Das gelegentlich vorkommende *pod volným nebem* ist stilistisch nicht angemessen (77 Google-Belege), die Wendung *pod otevřeným nebem* kommt häufig vor (780 Google-Belege). Es ist wohl dem Einfluss des engl. *in the open space* geschuldet.) Auch bei solch weitgehend äquivalenten Wortverbindungen muss natürlich gefragt werden, ob die Kontexte, in denen sie typischerweise vorkommen, ebenso äquivalent sind. Wie die Analyse von *frei* (*freier Himmel*) gezeigt hat (vgl. BELICA/STEYER hier S. 7), wird die Wortverbindung im Deutschen vor allem in zwei Kontexten verwendet: ‚Veranstaltungen/Feste/Gottesdienste‘ und ‚Übernachtung im Freien‘. Interessant ist an dieser Stelle, dass man das – natürlich auch im Deutschen sehr geläufige – englische *open air* nur für einige dieser Kontexte verwenden kann, für andere aber nicht. Veranstaltungen finden *unter freiem Himmel* = *open air* statt, aber eine Jugendgruppe würde zwar *unter freiem Himmel*, nicht aber *open air* übernachten, was auch die KA von *open air* bestätigt. Im Tschechischen gibt es eine typische feste Wendung, eine Entsprechung zu *unter freiem Himmel übernachten/schlafen*, und zwar in der Umgangssprache: *spát pod širákem* (‚unter einem breitkrempigen Hut schlafen‘; *širák* von *široký* abgeleitet). (*Weit*, *breit* sind Entsprechungen zu *široký*, *weit* auch als *širý* nur in bestimmten Wendungen vorhanden.)

- 2) Andere Beispiele seien nur aufgezählt: *einer Sache einen freien Lauf lassen* (po)nechat něčemu volný průběh, *freier Markt* volný trh, *Weg frei machen* uvolnit cestu; *Freie Theatergruppe* nezávislý divadelní soubor usw. Alle genannten Wortverbindungen sind im Wörterbuchartikel anzuführen, weil es dafür jeweils die passenden tschechischen Synonyme (volný, širý, nezávislý) gibt.
- 3) *frei Haus s donáškou do domu* – Hier lässt sich anhand der Kookkurrenzpartner erkennen, dass die Wortverbindung trotz syntaktischer Festigkeit polysem ist: Sie wird nicht nur in der Bedeutung ‚etwas direkt ins Haus bringen‘ (z.B. durch solche Kookkurrenzpartner wie *liefern, monatlich, Flaschen, Lieferungen* usw.) verwendet, sondern auch im übertragenden Sinne (‚etwas ohne großen Aufwand/Widerstand/„einfach so“ bekommen, erreichen‘ usw.). Entsprechende Kookkurrenzpartner sind beispielsweise: *Argumente, Wahlkampfmunition, quasi*. Die Wendung kann im Tschechischen z.B. in einer ironischen Bedeutung gebraucht werden, dieser Fall ist jedoch selten anzutreffen.

d. Vollständige Asymmetrien

Die KA hilft dem nicht-muttersprachlichen Anwender (in erster Instanz natürlich dem Lexikographen), die Bedeutungen und den typischen Verwendungskontext solcher nicht-transparenter, „äquivalentloser“ Wortverbindungen zu entschlüsseln und überhaupt erst zu verstehen.

- 1) *frei empfangbar* (*frei empfangbares Fernsehen* → *Sender* → *Programme*) (‚kein Privat- oder Bezahlfernsehen‘) nestátní, nezávislá televize. Kookkurrenzpartner sind u. a. *Fernsehen* → *Sender* → *Free-TV* → *Pay-TV* → *Programme* → *Bezahlfernsehen* → *Vox*.
- 2) *frank und frei* (‚offene und ehrliche Redeweise‘). Verbale Ergänzungen von *frank und frei* sind z.B. *erklären, seine Meinung sagen, zugeben, gestehen, bekennen, erzählen*. Die Kookkurrenzanalyse von *frank* ergibt eine ausgeprägte Gebundenheit an das Adjektiv *frei*. Man kann von einem unikalen Status sprechen, obwohl es einige wenige Belege in ungebundener Form gibt. Diese sind aber okkasionell und deshalb u. E. nicht in ein Wörterbuch wie das GDTAW aufzunehmen. Die Kookkurrenzanalyse bestätigt darüber hinaus die fixe Struktur der Zwillingsformel. Diese Festigkeit muss vermerkt werden. *frank* sollte daher auch nicht als Einzellexem übersetzt werden, sondern immer nur mit einem Äquivalent, dass die kommunikative Funktion der ganzen Wortverbindung angemessen wiedergibt.
- 3) *aus freien Stücken* (‚freiwillig, nicht unter Zwang‘). Diese Wortverbindung ist vollkommen fest, sowohl lexikalisch als auch strukturell hinsichtlich des phrasalen Status (Präpositionalphrase) als auch des Numerus (Plural), wie an dem Ausschnitt aus KWIC-Zeilen deutlich wird:

| | | |
|-----|--|---|
| T86 | Meldungen hieß es, die Bewohner hätten | aus freien Stücken einen neuen Wohnort gewählt. |
| T87 | daß einige Tataren die Hauptstadt | aus „freien Stücken“ verlassen hätten. |
| T88 | Kunden“ ohne Gummischutz, und das „nicht | aus freien Stücken, sondern weils eben Geld |
| T89 | Was bringt Zwölfjährige dazu, sich | aus freien Stücken vor Polizeiwannen zu setzen? |
| T89 | Wir haben uns | aus freien Stücken dazu entschieden. |
| T89 | handelt allerdings auch nicht | aus freien Stücken. |
| T89 | Und vor allem: tat sie dies | aus freien Stücken? |
| T89 | Er hat das Riesenreich im Osten | aus freien Stücken in Bewegung gebracht. |
| X99 | sich nicht an ihren Peinigern, sondern hat | aus freien Stücken Mitleid mit den Bösen. |
| X99 | dass er 1995 Washington und dem Senat | aus freien Stücken den Rücken gekehrt hat. |
| X00 | Wortmeldung hervorging, doch nicht so ganz | aus freien Stücken aus der Kommunalpolitik |
| X00 | Sie habe ihre Entscheidung | aus freien Stücken getroffen, da sie sich nicht |
| X00 | „Er ist | aus freien Stücken zur Polizei gegangen, um |
| M02 | und sogar einem Beil traktierte, nicht | aus freien Stücken mitgemacht: Er sei dazu |
| M02 | eines Verbrechens geworden, oder hat es | aus freien Stücken dem Elternhaus den Rücken |
| M02 | Thomas Jöhl trat | aus freien Stücken zurück und sein Nachfolger |
| M02 | „Wir haben ihn | aus freien Stücken aus der Taufe gehoben und |
| M03 | hatte sein Mandat zum 31. Dezember | aus freien Stücken niedergelegt. |
| M03 | und hofft, dass sich Bundesbürger | aus freien Stücken zwangskollektivieren, um |
| M03 | Mit 3,6 Promille Verkehr geregelt | Aus freien Stücken betätigte sich eine 37 Jahre |

Einzige Variationen bilden modifizierende Einschübe, die als Verstärker fungieren wie: *aus völlig/ganz... freien Stücken*. Die Invarianz dieser Wortverbindung ist eine wichtige Information für den Nutzer und daher im Wörterbuch zu vermerken.

e. Domänenbezogene Benennungen

In der KA findet sich eine Vielzahl von institutionellen Bezeichnungen mit dem Adjektiv *frei*, die auch für einen Wörterbuchnutzer von Interesse sind.

Freie Universität Berlin, Freie Demokraten, Freie Wählergemeinschaft, Bund Freier Bürger, Freie Evangelische Gemeinde usw.

Das GDTAW ist zugleich von der Arbeitsstätte geprägt, an der es entsteht. Weil das Institut seit Jahrzehnten auch als Beratungsstelle für andere Institute gilt, sind diese Ergebnisse noch im Hinblick auf diese Adressaten (z.B. Übersetzer) – auch unter Einbeziehung historischer Kontexte – zu vervollständigen (*Freie Hansestadt Bremen, die Freie Schulgemeinde Wickersdorf, Freie Deutsche Jugend, Freie Reichsstadt; Sie unterstellen sich als Freie Zinser einem Kloster. ... ; freie/unfreie Leihe* usw.) Diese Kollokationen sind dann im GDTAW bei den entsprechenden Substantiven auch zu lemmatisieren. (Die jeweiligen WBA werden vernetzt.)

Fazit

Der Einsatz der KA führt dazu, dass der Autor des Wörterbuchartikels seine auf eigener Introspektion und anderen Wörterbüchern gegründete Vorstellungen und Hypothesen revidieren und ergänzen kann. Er findet nicht nur neue, gut gebräuchliche Verwendungsmuster aus dem lebendigen Sprachgebrauch, die das Informationsangebot in einem Übersetzungswörterbuch erheblich aufwerten. Das größte Potenzial

bietet die KA für eine feine Ausdifferenzierung usueller Verwendungsmuster, für die Erfassung und Beschreibung eines sehr situativ gebundenen und kontextsensitiven Sprachgebrauchs „unterhalb“ der klassischen Lesarten, für stilistische Feinheiten, die aber besonders wichtige Kriterien für den Grad der Beherrschung einer Fremdsprache darstellen. Gerade hier liegt ja die große Herausforderung auch für professionelle Multiplikatoren des Deutschen wie Fremdsprachlehrer, Lexikografen und Übersetzer. Und hier liegen auch noch die Desiderata der bisherigen bi- und multilingualen Forschung, Fremdsprachdidaktik und Lexikografie. Statistisch ermittelte Kohäsionswerte zwischen sprachlichen Einheiten (z.B. der LLR-Wert) und Frequenzen sind in diesem Kontext sehr wichtige Anhaltspunkte für den Grad der Typikalität bzw. Usualität, können aus nicht-muttersprachlicher Sicht in letzter Instanz aber nur im Kontext anderer Quellen gesehen werden. Natürlich sollte man solche Methoden wie die statistische Kookkurrenzanalyse nicht als alleinige empirische Methoden ansehen. Grammatiken und Wörterbücher (z.B. das DUW) sind ebenso heranzuziehen wie bisherige sprachwissenschaftliche Ergebnisse sowie das Urteil von Sprechern mit muttersprachlicher Kompetenz. Auch wirft das in den Korpora und auch im Internet vorkommende „ungekämmte“ Wortmaterial in praktischer Hinsicht viele Fragen auf. Plädiert die moderne korpusbasierte Lexikografie für die Aufnahme authentischer Verwendungsstrukturen, dann ist Behutsamkeit geboten: Der auf kontrastive Sicht trainierte Wörterbuchschreiber gerät fast immer in die Lage, sich fragen zu müssen: Sollen im Wörterbuch wirklich *möglichst viele variable syntaktische Strukturen* dargestellt werden oder nur die, die als wirklich signifikant oder typisch erkannt wurden? Kommen wir dafür noch einmal auf die Wortverbindung *freier Verkauf* zurück. Wir hatten gezeigt, dass hier eine kontextabhängige Restriktion (Domäne ‚Verkauf von Eintrittskarten‘) in Hinblick auf ihre syntaktische Ausbaumöglichkeit vorliegt (vgl. S.): Die in diesem Verwendungskontext zu Recht als agrammatisch empfundene und durch die KA verifizierte Struktur **der Verkauf ist frei* wäre im GTDAW nicht zu lemmatisieren. Nun findet man aber mit Hilfe der Suchmaschine Google 123 Kontextstellen für syntaktisch erweiterte Konstruktionen in dem authentischen Satz *Der Verkauf ist frei ab 18 Jahren*. Auch in den IDS-Korpora ist diese Verwendungsweise belegt. 9 Belege bestätigen außerdem die Konstruktion im Satz *Der Verkauf ist ab 18 Jahren frei*. Sollte man diese syntaktischen Varianten trotz des im Vergleich zu den über 19 000 Vorkommen von *freier Verkauf* in den IDS-Korpora eher marginalen Vorkommens ebenfalls aufnehmen und wenn ja, inwieweit können und müssen diese Vorkommensverhältnisse und Kontextabhängigkeiten explizit kommentiert werden? Solche und ähnliche Fragen könnte man zwar mit dem Argument abtun, dass nur zentrale, und nicht periphere Erscheinungen in ein Wörterbuch aufzunehmen sind, vor allem, wenn es sich um ein passives Wörterbuch handelt. Gegen dieses Argument spricht aber die praktische Erfahrung der Wörterbuchbenutzer: Ein gutes Wörterbuch ist besser als jede Grammatik, die notwendigerweise verallgemeinern muss. Nur ein Wörterbuch kann auch punktuelle – eher als peripher anzusehende

– Fragen beantworten. Darüber hinaus bewegen solche Fragen die fortgeschrittenen tschechischen Übersetzer und Studenten, die das Wörterbuch als einen stilistischen Berater heranziehen. In unserem Beispiel der syntaktischen Ausbaumöglichkeiten von *freier Verkauf* und *freier Eintritt* helfen z.B. die einsprachigen deutschen Wörterbücher entweder gar nicht oder nur ausschnittsweise weiter. Die Konstruktion *frei – Verkauf* ist weder unter *Verkauf* noch unter *frei* zu finden. Bei *freier Eintritt* haben sich die Wörterbücher jeweils für eine der beiden syntaktischen Varianten entschieden: DUW führt nur die syntaktische Variante der attributiven Stellung *freier Eintritt* an. Als Beispiel für ein postponiertes Adjektiv *frei* ist dort *ein Getränk ist frei – jeden nápoj/jedno pití zdarma* verzeichnet. Im GWDS wird dagegen allein die prädikative Stellung *Der Eintritt ist frei* angegeben.⁵

Der vorliegende Artikel hatte das Ziel, den außerordentlichen Nutzen der Kookkurrenzanalyse für die kontrastive Betrachtung des Deutschen „von außen“ – wenn auch nur in einer kurzen Skizze – zu thematisieren. Ein weiteres Desiderat der kontrastiven Sprachforschung tritt jedoch bei diesem analytischen Vorgehen besonders deutlich hervor: die Notwendigkeit einer parallelen Anwendung solcher statistischen Methoden wie der Kookkurrenzanalyse in den Korpora der beteiligten Sprachen, in unserem Falle also in einem in einem deutschen **und** in einem tschechischen Korpus. Erst dann könnte man wirklich vergleichend ermitteln, wie sich Kookkurrenzprofile einzelner Lexeme oder auch Lexemgruppen zueinander verhalten, welche Überschneidungen und Abweichungen es gibt. Dies würde dann in der Tat eine völlig Neuorientierung in methodischer Hinsicht zur Folge haben.

Literatur

AG für Korpustechnologie (2004) (<http://www.ids-mannheim.de/kt/>): Eine kurze Einführung in die Kookkurrenzanalyse und syntagmatische Muster unter <http://www.ids-mannheim.de/kt/misc/tutorial.html>

BAHNS, J. (1997): *Kollokationen und Wortschatzarbeit im Englischunterricht*. Gunter Narr Verlag, Tübingen

BURGER, H. (1998): *Phraseologie. Eine Einführung am Beispiel des Deutschen*. Berlin, Schmidt.

BELICA, C. – STEYER, K.: Korpusanalytische Zugänge zu sprachlichem Usus (in diesem Band).

HAUSMANN, F. J. (1984): Wortschatzlernen ist Kollokationslernen. Zum Lehren und Ler-

⁵ Das Problem der unterschiedlichen Beschaffenheit des deutschen und tschechischen Adjektivs wird in diesem Sammelband noch in einem weiteren Aufsatz aufgegriffen.

- nen französischer Wortverbindungen. In: *Praxis des neusprachlichen Unterrichts*, Jg. 31, 395–406
- (1985): Kollokationen im deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels. In: Bergenholtz, H. – Mugdan, J. (Hg.), *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch*, Lexicographica 3, Tübingen, 118–129
- (1989): Le dictionnaire de collocations. Artikel 95, in: *Wörterbücher. Dictionaries. Dictionnaires*. Hrsg. v. F. J. Hausmann – O. Reichmann – H. E. Wiegand – L. Zgusta. Erster Teilband. Berlin 1989, 1010–1019.
- (2004): Was sind eigentlich Kollokationen? In: Steyer, Kathrin (Hg.), *Wortverbindungen – mehr oder weniger fest*, Berlin/New York (= Jahrbücher des Instituts für Deutsche Sprache, 2003), 309–334
- KUBCZAK, J. – COSTANTINO, S. (1998): in: Bresson, D. – Kubczak, J. (Hg.), *Abstrakte Nomina. Vorarbeiten zu ihrer Erfassung in einem zweisprachigen syntagmatischen Wörterbuch* (= Studie zur deutschen Sprache 10), Tübingen, 11–120
- PERKUHN, R. – BELICA, C. – AL WADI, D. – LAUER, M. – STEYER, K. – WEISS, CH. (2005): Korpustechnologie am Institut für Deutsche Sprache. In: *Tagungsband der Internationalen Konferenz Korpuslinguistik deutsch: synchron – diachron – kontrastiv*; 20.–23. 3. 2003, Universität Würzburg, 57–70
- STEYER, K. (1998): Kollokationen als zentrales Übersetzungsproblem – Vorschläge für eine Kollokationsdatenbank Deutsch-Französisch/Französisch-Deutsch auf der Basis paralleler und vergleichbarer Korpora. In: Bresson, D. (Hrsg.), *Lexikologie und Lexikographie Deutsch-Französisch*, Aix-en-Provence (= Cahiers d' Études Germaniques, 35), 95–113.
- (2000): Usuelle Wortverbindungen des Deutschen. Linguistisches Konzept und lexikografische Möglichkeiten. In: *Deutsche Sprache* 2/00, 101–125.
- (2004b): Kookkurrenz. Korpusmethodik, linguistisches Modell, lexikografische Perspektiven. In: Steyer, K. (Hrsg.): *Wortverbindungen – mehr oder weniger fest*. Berlin/New York (= Jahrbücher des Instituts für Deutsche Sprache, 2003), 87–116.
- VACHKOVÁ, M. (2002): Das große deutsch-tschechische Wörterbuchprojekt im Kontext der zeitgenössischen deutsch-tschechischen Übersetzungslexikographie. In: Kunzmann-Müller, B., Sprachwandel und Lexikographie / M. Zielinski (Hrsg.): *Beispiele aus slawischen Sprachen, dem Ungarischen und Albanischen* – Frankfurt am Main [u. a.]: Lang, 2002. (Berliner slawistische Arbeiten)
- (2003): Der deutsch-tschechische Sprachvergleich im deutsch-tschechischen Übersetzungswörterbuch. In: *brücken. Germanistisches Jahrbuch Tschechien–Slowakei*, Neue Folge 11, 251–264.

Wörterbücher

Siebenschein, H.: *Německo-český slovník. Deutsch-tschechisches Wörterbuch* (2002): I. A–L II. M–, Leda, 7., überarb. Auflage, Praha, 2 Bde.

Duden – Das große Wörterbuch der deutschen Sprache (2000), hrsg. vom Wissenschaftlichen Rat der Dudenredaktion. (CD-ROM; basiert auf der 3., völlig neu bearbeiteten und erweiterten Auflage der Buchausgabe in 10 Bänden – 1999.)

Duden – Deutsches Universalwörterbuch (2003): 5. Aufl. Mannheim (CD-ROM)

COOCCURRENCE ANALYSIS SEEN CONTRASTIVELY

On applying collocational patterning in bilingual lexicography – some examples from the large German-Czech academic dictionary

Summary

This paper resumes some of thoughts presented in the study by C. Belica and K. Steyer in this volume. It shows how bilingual lexicographers can take advantage of the cooccurrence analysis results when dealing with German-Czech contrast and structuring word configurations in an entry. They also sketch the corpus data in a form of structural types based on the collocational patterns and stress the importance of cooccurrence analysis for an enlarged offer of equivalents. They plead for more consideration of the syntactic variability. They argue that the cooccurrence analysis used for both German and for Czech should be an important step.

KOOCURENČNÍ ANALÝZA V KONTRASTIVNÍM POHLEDU

O jejím využití při vzniku velkého německo-českého akademického slovníku

Resumé

Článek navazující na příspěvek autorů C. Belici a K. Steyerové v tomto sborníku ukazuje užitečnost kookurenční analýzy pro kontrastivní pohled na německou a českou slovní zásobu při práci na heslových statích. Možnosti jeho zmapování dokumentují autorky na nástinu několika charakteristických strukturních celků. Vedle podchycení často se vyskytujících slovních spojení a významné pomoci při vytváření nabídky překladových ekvivalentů se otevírají i nové možnosti pro zohlednění syntaktických obměn. Autorky se přimlouvají za uplatnění kookurenční analýzy při výzkumu obou jazyků, nejen němčiny.